REPORT DOCUMENTATION PAGE

AFRL-SR-BL-TR-98-

0709

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	9/1/1998	Final Technical Report 3/1/94 to 9/30/97

4. TITLE AND SUBTITLE
Early Scene Analysis: Rapid Processing of Contours, Surfaces, and Objects in Human Vision

5. FUNDING NUMBERS
AFOSR 94-0189
2313/BS
61102F

6. AUTHOR(S)
Patrick Cavanagh

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Harvard University
1350 Massachusetts Avenue
Holyoke Center
Cambridge, MA 02138-3826

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR / NL
Life Science Directorate
Bldg 410
Bolling AFB, DC 20332-6448

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release;
distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT ***(Maximum 200 Words)***

How does the human brain represent objects? How does it recognize them so rapidly? We have been able to show how 2-D information is built up from the parallel analysis of a set of visual attributes and how this information contacts memory in order to construct 3-D representations of the visual scene. We have demonstrated a simplified early contact with memory, occuring prior to any part-based or contour analysis an yet capable of guiding recognition. We have described a motion phenomenon which reveals the nature of image segmentation in a very direct manner. We have discovered a new technique to probe the dimensions of high-level shape recognition and we have described a new level of object description in terms of volumes which subsumes earlier work on contours and surfaces. Finally, work on shading and shadows is contributing to a new approach to rapid line labeling in images.

14. SUBJECT TERMS	15. NUMBER OF PAGES
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
U	U	U	UL

EARLY SCENE ANALYSIS: RAPID PROCESSING OF CONTOURS, SURFACES, AND OBJECTS IN HUMAN VISION

Summary

How does the human brain represent objects? How does it recognize them so rapidly? We have been able to show how 2-D information is built up from the parallel analysis of a set of visual attributes and how this information contacts memory in order to construct 3-D representations of the visual scene. We have demonstrated a simplified early contact with memory, occuring prior to any part-based or contour analysis an yet capable of guiding recognition. We have described a motion phenomenon which reveals the nature of image segmentation in a very direct manner. We have discovered a new technique to probe the dimensions of high-level shape recognition and we have described a new level of object description in terms of volumes which subsumes earlier work on contours and surfaces. Finally, work on shading and shadows is contributing to a new approach to rapid line labeling in images.

Status of Effort

All the proposed projects of the grant got underway, several were completed and a few are continuing along with new projects into the next grant. Overall, 13 articles have been published based on work funded by this grant. Details are given in the next section.

Accomplishments / New Findings

Cross-media cooperation in contour localization. Contour localization appears to be based on contributions from all attributes. Luminance does not play a special or predominant role (Rivest & Cavanagh, 1995). No further work is planned.

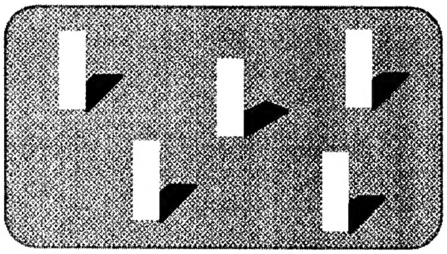
Cross-media pictorial cues. Surface slant is available from images whether depicted in luminance or in color (Zimmerman, Legge & Cavanagh, 1995). No further work is planned.

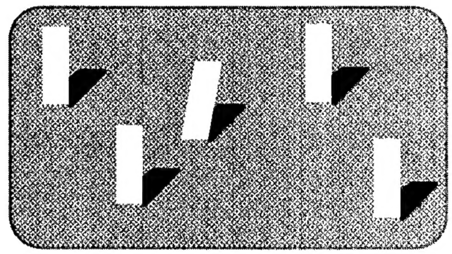
Access to features in visual search. When search stimuli form high-level patterns such as faces, search based on direct access to low-level feature (eg the arcs and strokes of the eyes or mouth) is no longer possible even though it would be faster. High-level context can therefore slow search in some cases (Suzuki & Cavanagh, 1995). No further work is planned.

Familiarity as a feature in visual search. Search for an unfamiliar pattern in a field of familiar patterns was found to be very rapid (parallel) while the reverse was not (serial). This rapid search was not based on any low-level feature but solely on the familiarity of the pattern. Novelty seems to attract attention automatically. This seems reasonable as novel items are in general the ones requiring additional processing (Wang, Cavanagh, & Green, 1994). No further work is planned.

Object features and scene attributes in visual search. During the previous grant period, Ron Rensink and I showed that preattentive vision is sensitive to scene structure defined by shadows and highlights (Rensink & Cavanagh, 1994). We have developed a more general paradigm where the subject reports the odd item which may

differ from the distractors with respect to an illumination feature (shadow or highlight orientation) or an object feature (object orientation). We hope to discover the stage at which illumination features are identified and discounted relative to the stage at which visual search operates. This research is continuing.





Odd item is shadow Odd item is object

Analysis of shadows and shading. Finally, our interest in the internal representation of shading focused on the interpretation of shading images shown in reverse contrast. We have previously shown the expected importance of contrast polatity for shadows — when shadowed images are presented in reverse contrast, 3-D structure is often distorted or lost. This does not appear to be the case for shaded images. We captured objects and live scenes in our laboratory using a camera with the lighting fixed to camera lens. This lighting produces only shading in the image with no cast shadows. The only shadows arise from secondary reflection which was minimized with dark walls and materials. When viewed in negative contrast, these images appeared as convincingly 3-D as the positive versions. Our conclusion is that our visual system recovers depth from shading locally using brightness gradient cues and does not attempt to determine a common direction of lighting for the scene. This work was presented to ARVO (Cavanagh, 1995). This work continues into the next grant as the basis of an algorithmic approach to line labeling.

Position distortions. Satoru Suzuki and I discovered that briefly presented tests undergo significant distortions in apparent position and shape. The position distortions are found when an attention grabbing cue is followed by a brief test probe. The apparent position of the test is displaced away from the attention cue by up to 30 minutes of arc. This distortion may be a result of the recruitment of resources (migration of receptive fields) toward the attentional focus. Population coding models of location predict this same displacement when receptive fields migrate toward the attentional locus. This work was presented at ARVO (Suzuki & Cavanagh, 1994) and was published in *JEP:HPP*. No further work is planned.

Shape distortions. The shape of a briefly presented test can be influenced by the shape of a preceding cue. We have used this to develop a completely new paradigm which can catalog the dimensions of shape. A briefly presented line target will make a subsequent circle appear elliptical with the major axis orthogonal to the orientation of the line. This distortion appears to be global — its effect is largely indepenent of the offset between the line and the circle. Equivalent interactions are found between curved lines and straight lines and between trapezoids and squares. Our current hypothesis is that the distortion arises in shape-specific units (perhaps in inferotemporal cortex) which mutually inhibit each other. This work has been presented at ARVO (Suzuki & Cavanagh, 1995) and has just appeared in*JEP:HPP*.

Object recovery from 2 tone images. Dr. Moore studied simple and complex objects depicted in 2 tone images. She imaged single, simple shapes with direct lighting

that produces sharp cast shadows. Her first observation was that no single, simple shape is recognizable in a two-tone format (high-contrast, black and white). Complex objects like faces are fully interpretable in 2 tone versions but the very same parts rearranged into an unfamiliar shape loses its three-dimensionality when presented as a 2 tone image. Evidently, familiarity is a requirement for interpretation, arguing strongly against any part-based approach. This study has just appeared in *Cognition*. No further work is planned.

Priming 2 tone images with gray-scale images. Dr. Moore also began a study of the effects of a preview of a gray-scale version on the interpretation of a 2-tone image. Without the preview, the 2-tone images were seldom seen as the original object. The preview prime was the same object in the same or different view, or the same or different lighting. The goal was to determine whether the internal representation which was being primed was object centered or viewer centered and whether it was illumination dependent. Several classes of objects were used: familiar objects, unfamiliar objects, simple and mulitpart objects. This work continues.

Object recognition: positive priming. In our model, recognition starts with an initial, crude 2-D match that selects a "best" prototype to explain the image data. This is followed by more sophisticated 3-D analyses to complete the recognition process. Our first experiment showed a priming effect of contours in recognition even though the contours alone were uniformative for the task. This work was presented at ARVO (Cavanagh & Watanabe, 1996) and continues into the next grant with the collaboration of David Whitney a new graduate student.

Object models and motion perception. Peter Tse has developed a new theory for apparent motion that relies on parsing each scene into objects before matching takes place. The novel aspect of the work is that the shapes in the first frame of the motion sequence overlap spatially with those in the following frame. This enables Peter to test for principles of shape parsing (continuity, surface similarity, contiguity) that do not come into play in standard apparent motion where the shapes do not overlap. This give us a new tool for understanding image segmentation. This work was presented at ARVO (Tse, Cavanagh, & Nakayama, 1995, 1996) and is published in a recent book (Tse, Cavanagh, & Nakayama, 1998).

Theory of volume. Peter Tse has developed a new theory for the level at which objects are represented in understanding visual scenes. This work is exceptionally novel and important. Rather than depending on relations between image contours or inferring surfaces, Peter shows that the underlying mode of representation is one of volumes or occupied space. Several critical demonstrations show that his formulation accounts for the broad range of image interpretations whereas representations of objects by their contours or surfaces fail. Peter has three papers published or in press on this topic (Tse & Albert, 1998; Tse, 1998a, 1998b) and has present one talk (Tse, 1997)

Personnel supported

Personnel on the grant were myself (50% summer salary), Ron Rensink (Postdoctoral Fellow, 1994-95), Cassandra Moore (Postdoctoral Fellow, 1995-97), Satoru Suzuki (graduate student and summer research assistant, 1994-96), Raynald Comtois, our Senior Systems Analyst (25% salary), and Seth Hamlin, our research assistant (25% salary). Dr. Sheng He and Dr. Marvin Chun were both briefly salaried as postdoctoral fellows on this grant before being awarded NRSA Postdoctoral Fellowships.

Publications supported by grant

Wang, Q., Cavanagh, P., & Green, M. (1994). Familiarity and pop-out in visual search. *Perception & Psychophysics*, **56**, 495-500.

Rivest, J., & Cavanagh, P. (1995). Localizing contours defined by more than one attribute. *Vision Research*, **36**, 53-66.

Suzuki, S., & Cavanagh, P. (1995). Facial organization blocks access to low-level features: an object inferiority effect. *Journal of Experimental Psychology: Human Perception and Performance*, **21**, 901-913.

Zimmerman, D. L., Legge, G. E., & Cavanagh, P. (1995). Pictorial depth cues: A new slant. *Journal of the Optical Society of America A*, **12**, 17-26.

Cavanagh, P. (1996). Vision is getting easier every day. *Perception*, **24**, 1227-1232.

Watanabe, T., & Cavanagh, P. (1996). Texture laciness. *Perception*, **25**, 293-304.

Suzuki, S., & Cavanagh, P. (1997). Focused attention distorts visual space: An attentional repulsion effect. *Journal of Experimental Psychology: Human Perception and Performance*, **23**, 443-463.

Moore, C., & Cavanagh, P. (1998). Recovery of 3D volume from 2-tone images of novel objects. *Cognition*, **67**, 45-71.

Suzuki, S., & Cavanagh, P. (1998). A shape-contrast effect for briefly presented stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, **24**, 1315-1341.

Tse, P., Cavanagh, P., & Nakayama, K. (1998). The role of parsing in high-level motion processing. In Takeo Watanabe (ed.), *High level motion processing.* (pp. 249-266), Cambridge, MA: MIT Press.

Tse, P. & Albert, M. (1998). Amodal completion in absence of image tangent discontinuities. *Perception*, **27**, 455-464.

Tse, P. (1998a). Illusory volumes from conformation. *Perception*, in press.

Tse, P. (1998b). Volume completion. *Cognitive Psychology*, in press.

Interactions, conference papers supported by grant

Rensink, R., & Cavanagh, P. (1994). Identification of highlights in early vision. *Investigative Ophthalmology and Visual Science*, **35**, 1623.

Rivest, J., & Cavanagh, P. (1994). Spatial interactions between color contours and between luminance contours. *Investigative Ophthalmology and Visual Science*, **35**, 2008.

Suzuki, S., & Cavanagh, P. (1994). Focused attention distorts visual space. *Investigative Ophthalmology and Visual Science*, **35**, 2081.

Cavanagh, P. (1995). A horse of a different color: Shadows have to be darker but shading does not. *Investigative Ophthalmology & Visual Science*, **36**, S184.

Suzuki, S., & Cavanagh, P. (1995). A rapid sequence of two stimuli reveals nonretinotopic shape distortions. *Investigative Ophthalmology & Visual Science*, **36**, S848.

Tse, P., Cavanagh, P. & Nakayama, K. (1995). Line motion occurs after surface parsing. *Investigative Ophthalmology & Visual Science*, **36**, S417.

Cavanagh, P. (1996). Research in visual perception: The 40th millennium. *Journal Of International Psychology*, **31**, 530.1.

Cavanagh, P., & Watanabe, T. (1996). Priming object recognition with unrecognizable outlines. *Investigative Ophthalmology & Visual Science*, **37**, S193.

Moore, C., & Cavanagh, P. (1996). Discriminating shadow and object regions in 2 tone images. *Investigative Ophthalmology & Visual Science*, **37**, S193.

Tse, P., Cavanagh, P., & Nakayama, K. (1996). The roles of attention in shape change apparent motion. *Investigative Ophthalmology & Visual Science*, **37**, S213.

Tse, P., Intriligator, J., & Cavanagh, P. (1997). Attention distorts the perception of time. *Investigative Ophthalmology & Visual Science*, **38**, S1151. (A)

Tse, P. (1997). Plasmas: A new class of motion-induced brightness illusions. ECVP, Helsinki.

New discoveries, inventions, or patent disclosures

Discoveries are reported in the progress section above. There were no inventions or patent disclosures during the grant period.

Honors / Awards

None.